

0570  
07/16

ENTERED

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/067,974

DATE: 04/25/2002 P16  
TIME: 12:21:53

Input Set : A:\seqlist-15332640001 ascii

Output Set: N:\CRF3\04252002\J067974.raw

```

5 <110> APPLICANT: Li, Lhing-Yew
6   Trei, Kelli J.
9 <120> TITLE OF INVENTION: Polynucleotide Constructs for Increased Lysine Production
13 <130> FILE REFERENCE: 1533.2640001
C--> 17 <140> CURRENT APPLICATION NUMBER: US/10/067,974
C--> 17 <141> CURRENT FILING DATE: 2002-02-08
17 <150> PRIOR APPLICATION NUMBER: 60/267,183
19 <151> PRIOR FILING DATE: 2001-02-08
23 <160> NUMBER OF SEQ ID NOS: 25
27 <170> SOFTWARE: PatentIn version 3.1
32 <210> SEQ ID NO: 1
34 <211> LENGTH: 1266
36 <212> TYPE: DNA
38 <213> ORGANISM: Corynebacterium glutamicum
42 <220> FEATURE:
44 <221> NAME/KEY: CDS
46 <222> LOCATION: (1)..(1266)
50 <400> SEQUENCE: 1
51 gtg gcc ctg gtc gta cag aaa tat ggc ggt tcc tcg ctt gag agt gcg      48
52 Met Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala
53 1          5          10          15
55 gaa cgc att aga aac gtc gct gaa cgg atc gtt gcc acc aag agt gct      96
56 Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala
57          20          25          30
59 gga aat gat gtc gtg gtt gtc tgc tcc gca atg gga gac acc acg gat      144
60 Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp
61          35          40          45
63 gaa ctt cta gaa ctt gca gcg gca gtg aat ccc gtt ccg cca gct cgt      192
64 Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg
65          50          55          60
67 gaa atg gat atg ctc ctg act gct ggt gag cgt att tct aac gct ctc      240
68 Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu
69 65          70          75          80
71 gtc gcc atg gct att gag tcc ctt ggc gca gaa gct caa tct ttc act      288
72 Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr
73          85          90          95
75 ggc tct cag gct ggt gtg ctc acc acc gag cgc cac gga aac gca cgc      336
76 Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg
77          100          105          110
79 att gtt gac gtc aca ccg ggt cgt gtg cgt gaa gca ctc gat gag ggc      384
80 Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly
81          115          120          125
83 aag atc tgc att gtt gct ggt ttt cag ggt gtt aat aaa gaa acc cgc      432

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/067,974

DATE: 04/25/2002

TIME: 12:21:53

Input Set : A:\seqlist-15332640001 ascii

Output Set: N:\CRF3\04252002\J067974.raw

```

84 Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg
85      130      135      140
87 gat gtc acc acg ttg ggt cgt ggt ggt tct gac acc act gca gtt gcg      480
88 Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala
89 145      150      155      160
91 ttg gca gct gct ttg aac gct gat gtg tgt gag att tac tcg gac gtt      528
92 Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val
93      165      170      175
95 gac ggt gtg tat acc gct gac ccg cgc atc gtt cct aat gca cag aag      576
96 Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys
97      180      185      190
99 ctg gaa aag ctc agc ttc gaa gaa atg ctg gaa ctt gct gct gtt ggc      624
100 Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly
101      195      200      205
103 tcc aag att ttg gtg ctg cgc agt gtt gaa tac gct cgt gca ttc aat      672
104 Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn
105      210      215      220
107 gtg cca ctt cgc gta cgc tcg tct tat agt aat gat ccc ggc act ttg      720
108 Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu
109 225      230      235      240
111 att gcc ggc tct atg gag gat att cct gtg gaa gaa gca gtc ctt acc      768
112 Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr
113      245      250      255
115 ggt gtc gca acc gac aag tcc gaa gcc aaa gta acc gtt ctg ggt att      816
116 Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile
117      260      265      270
119 tcc gat aag cca ggc gag gct gcc aag gtt ttc cgt gcg ttg gct gat      864
120 Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp
121      275      280      285
123 gca gaa atc aac att gac atg gtt ctg cag aac gtc tcc tct gtg gaa      912
124 Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu
125      290      295      300
127 gac ggc acc acc gac atc acg ttc acc tgc cct cgc gct gac gga cgc      960
128 Asp Gly Thr Thr Asp Ile Thr Phe Thr Cys Pro Arg Ala Asp Gly Arg
129 305      310      315      320
131 cgt gcg atg gag atc ttg aag aag ctt cag gtt cag ggc aac tgg acc      1008
132 Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr
133      325      330      335
135 aat gtg ctt tac gac gac cag gtc ggc aaa gtc tcc ctc gtg ggt gct      1056
136 Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala
137      340      345      350
139 ggc atg aag tct cac cca ggt gtt acc gca gag ttc atg gaa gct ctg      1104
140 Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu
141      355      360      365
143 cgc gat gtc aac gtg aac atc gaa ttg att tcc atc tct gag atc cgc      1152
144 Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Ile Ser Glu Ile Arg
145      370      375      380
147 att tcc gtg ctg atc cgt gaa gat gat ctg gat gct gct gca cgt gca      1200
148 Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/067,974

DATE: 04/25/2002

TIME: 12:21:53

Input Set : A:\seqlist-15332640001 ascii

Output Set: N:\CRF3\04252002\J067974.raw

```

149 385          390          395          400
151 ttg cat gag cag ttc cag ctg ggc ggc gaa gac gaa gcc gtc gtt tat      1248
152 Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr
153          405          410          415
155 gca ggc acc gga cgc taa      1266
156 Ala Gly Thr Gly Arg
157          420
160 <210> SEQ ID NO: 2
162 <211> LENGTH: 421
164 <212> TYPE: PRT
166 <213> ORGANISM: Corynebacterium glutamicum
170 <400> SEQUENCE: 2
172 Met Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala
173 1          5          10          15
176 Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala
177          20          25          30
180 Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp
181          35          40          45
184 Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg
185          50          55          60
188 Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu
189 65          70          75          80
192 Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr
193          85          90          95
196 Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg
197          100          105          110
200 Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly
201          115          120          125
204 Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg
205          130          135          140
208 Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala
209 145          150          155          160
212 Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val
213          165          170          175
216 Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys
217          180          185          190
220 Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly
221          195          200          205
224 Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn
225          210          215          220
228 Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu
229 225          230          235          240
232 Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr
233          245          250          255
236 Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile
237          260          265          270
240 Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp
241          275          280          285
244 Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu

```

## RAW SEQUENCE LISTING

DATE: 04/25/2002

PATENT APPLICATION: US/10/067,974

TIME: 12:21:53

Input Set : A:\seqlist-15332640001 ascii

Output Set: N:\CRF3\04252002\J067974.raw

```

245      290      295      300
248 Asp Gly Thr Thr Asp Ile Thr Phe Thr Cys Pro Arg Ala Asp Gly Arg
249 305      310      315      320
252 Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr
253      325      330      335
256 Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala
257      340      345      350
260 Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu
261      355      360      365
264 Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Ile Ser Glu Ile Arg
265      370      375      380
268 Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala
269 385      390      395      400
272 Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr
273      405      410      415
276 Ala Gly Thr Gly Arg
277      420
280 <210> SEQ ID NO: 3
282 <211> LENGTH: 1035
284 <212> TYPE: DNA
286 <213> ORGANISM: Corynebacterium glutamicum
290 <220> FEATURE:
292 <221> NAME/KEY: CDS
294 <222> LOCATION: (1)..(1035)
298 <400> SEQUENCE: 3
299 atg acc acc atc gca gtt gtt ggt gca acc ggc cag gtc ggc cag gtt      48
300 Met Thr Thr Ile Ala Val Val Gly Ala Thr Gly Gln Val Gly Gln Val
301 1      5      10      15
303 atg cgc acc ttt ttg gaa gag cgc aat ttc cca gct gac act gtt cgt      96
304 Met Arg Thr Phe Leu Glu Glu Arg Asn Phe Pro Ala Asp Thr Val Arg
305      20      25      30
307 ttc ttt gct tcc ccg cgt tcc gca ggc cgt aag att gaa ttc cgt ggc      144
308 Phe Phe Ala Ser Pro Arg Ser Ala Gly Arg Lys Ile Glu Phe Arg Gly
309      35      40      45
311 acg gaa atc gag gta gaa gac att act cag gca acc gag gag tcc ctc      192
312 Thr Glu Ile Glu Val Glu Asp Ile Thr Gln Ala Thr Glu Glu Ser Leu
313      50      55      60
315 aag ggc atc gac gtt gcg ttg ttc tct gct gga ggc acc gct tcc aag      240
316 Lys Gly Ile Asp Val Ala Leu Phe Ser Ala Gly Gly Thr Ala Ser Lys
317 65      70      75      80
319 cag tac got cca ctg ttt gct gct gca ggc gcg act gtt gtg gat aac      288
320 Gln Tyr Ala Pro Leu Phe Ala Ala Ala Gly Ala Thr Val Val Asp Asn
321      85      90      95
323 tct tct gct tgg cgc aag gac gac gag gtt cca cta atc gtc tct gag      336
324 Ser Ser Ala Trp Arg Lys Asp Asp Glu Val Pro Leu Ile Val Ser Glu
325      100      105      110
327 gtg aac cct tcc gac aag gat tcc ctg gtc aag ggc att att gcg aat      384
328 Val Asn Pro Ser Asp Lys Asp Ser Leu Val Lys Gly Ile Ile Ala Asn
329      115      120      125

```

## RAW SEQUENCE LISTING

DATE: 04/25/2002

PATENT APPLICATION: US/10/067,974

TIME: 12:21:53

Input Set : A:\seqlist-15332640001 ascii

Output Set: N:\CRF3\04252002\J067974.raw

```

331 cct aac tgc acc acc atg gct gca atg cca gtg ctg aag cca ctg cac      432
332 Pro Asn Cys Thr Thr Met Ala Ala Met Pro Val Leu Lys Pro Leu His
333      130      135      140
335 gat gcc gct ggt ctt gta aag ctt cac gtt tcc tct tac cag gct gtt      480
336 Asp Ala Ala Gly Leu Val Lys Leu His Val Ser Ser Tyr Gln Ala Val
337 145      150      155      160
339 tcc ggt tct ggt ctt gca ggt gtg gaa acc ttg gca aag cag gtt gct      528
340 Ser Gly Ser Gly Leu Ala Gly Val Glu Thr Leu Ala Lys Gln Val Ala
341      165      170      175
343 gca gtt ggc gac cac aac gtt gag ttc gtc cat gat gga cag gct gct      576
344 Ala Val Gly Asp His Asn Val Glu Phe Val His Asp Gly Gln Ala Ala
345      180      185      190
347 gac gca ggc gat gtc gga cct tac gtt tcc cca atc gct tac aac gtg      624
348 Asp Ala Gly Asp Val Gly Pro Tyr Val Ser Pro Ile Ala Tyr Asn Val
349      195      200      205
351 ctg cca ttc gcc gga aac ctc gtc gat gac ggc acc ttc gaa acc gac      672
352 Leu Pro Phe Ala Gly Asn Leu Val Asp Asp Gly Thr Phe Glu Thr Asp
353      210      215      220
355 gaa gag cag aag ctg cgc aac gaa tcc cgc aag att ctc ggc ctc cca      720
356 Glu Glu Gln Lys Leu Arg Asn Glu Ser Arg Lys Ile Leu Gly Leu Pro
357 225      230      235      240
359 gac ctc aag gtc tca ggc acc tgc gtc cgc gtg ccg gtt ttc acc ggc      768
360 Asp Leu Lys Val Ser Gly Thr Cys Val Arg Val Pro Val Phe Thr Gly
361      245      250      255
363 cac acg ctg acc att cac gcc gaa ttc gac aag gca atc acc gtc gag      816
364 His Thr Leu Thr Ile His Ala Glu Phe Asp Lys Ala Ile Thr Val Glu
365      260      265      270
367 cag gcg cag gag atc ttg ggt gcc gct tca ggc gtc gag ctt gtc gac      864
368 Gln Ala Gln Glu Ile Leu Gly Ala Ala Ser Gly Val Glu Leu Val Asp
369      275      280      285
371 gtc cca acc cca ctt gca gct gcc ggc att gac gaa tcc ctc gtt gga      912
372 Val Pro Thr Pro Leu Ala Ala Ala Gly Ile Asp Glu Ser Leu Val Gly
373      290      295      300
375 cgc atc cgt cag gac tcc act gtc gac gac aac cgc ggt ctg gtt ctc      960
376 Arg Ile Arg Gln Asp Ser Thr Val Asp Asp Asn Arg Gly Leu Val Leu
377 305      310      315      320
379 gtc gta tct ggc gat aac ctt cgc aag ggc gca gca ctg aac acc att      1008
380 Val Val Ser Gly Asp Asn Leu Arg Lys Gly Ala Ala Leu Asn Thr Ile
381      325      330      335
383 cag att gct gag ctg ctg gtt aag taa      1035
384 Gln Ile Ala Glu Leu Leu Val Lys
385      340
388 <210> SEQ ID NO: 4
390 <211> LENGTH: 344
392 <212> TYPE: PRT
394 <213> ORGANISM: Corynebacterium glutamicum
398 <400> SEQUENCE: 4
400 Met Thr Thr Ile Ala Val Val Gly Ala Thr Gly Gln Val Gly Gln Val
401 1      5      10      15

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/067,974

DATE: 04/25/2002  
TIME: 12:21:54

Input Set : A:\seqlist-15332640001 ascii  
Output Set: N:\CRF3\04252002\J067974.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:16; Xaa Pos. 40,317,345,380

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:17,18,19,20,21,22,23,24,25

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/067,974

DATE: 04/25/2002

TIME: 12:21:54

Input Set : A:\seqlist-15332640001 ascii

Output Set: N:\CRF3\04252002\J067974.raw

L:17 M:270 C: Current Application Number differs, Replaced Current Application No  
L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:1472 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:32  
L:1540 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:304  
L:1548 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:336  
L:1556 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:368  
L:1580 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17  
L:1598 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:18  
L:1616 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:19  
L:1634 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:20  
L:1652 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:21  
L:1670 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22  
L:1690 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23  
L:1708 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:24  
L:1727 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:25